

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: WEI et al

Serial No. To Be Assigned

Filed: February 16, 2001

**For: ISOLATED HUMAN DRUG-METABOLIZING
PROTEINS, NUCLEIC ACID MOLECULES
ENCODING HUMAN DRUG-METABOLIZING
PROTEINS, AND USES THEREOF**

Art Unit:

Examiner:

Atty. Docket: CL000763

**SUBMISSION OF SEQUENCE LISTING
UNDER 37 C.F.R. § 1.821(a)**

Honorable Commissioner of
Patents and Trademarks
Washington, D.C. 20231

Sir:

In compliance with 37 C.F.R. § 1.821(a), Applicants submit the Sequence Listing,
including the paper copy of the Sequence Listing and the computer readable copy of the
Sequence Listing.

In the Specification:

Please enter the Sequence Listing between the specification and the claims of the
above-identified application.

REMARKS

In accordance with 37 C.F.R. § 1.821(f), the paper copy of the Sequence Listing and the computer readable copy of the Sequence Listing submitted herewith in the above application are the same.

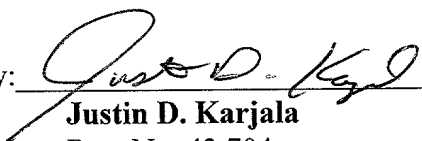
In accordance with 37 C.F.R. § 1.821(g), this submission includes no new matter.

It is respectfully believed this application is now in condition for examination.

Early notice to this effect is earnestly solicited.

Respectfully submitted,

CELERA GENOMICS

By: 
Justin D. Karjala
Reg. No. 43,704

Date: February 16, 2001

Celera Genomics Corporation
45 West Gude Drive, C2-4#20
Rockville, MD 20850
Tel: 240-453-3067
Fax: 240-453-3084

SEQUENCE LISTING

<110> WEI, Ming-Hui et al.

<120> ISOLATED HUMAN DRUG-METABOLIZING
PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
DRUG-METABOLIZING PROTEINS,
AND USES THEREOF

<130> CL000763

<160> 5

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 2759

<212> DNA

<213> Human

<400> 1

```

caaccattgc agatcagtgt gtgagggaac tgccatcatg aggtctgaca agtcagcttt 60
ggtattttctg ctccctgcagc tcttctgtgt tggctgtgga ttctgtggga aagtcctggt 120
gtggccctgt gacatgagcc attggcttaa tgtcaaggtc attctagaag agtcatagt 180
gagaggccat gaggtaacag tattgactca ctcaaagcct tcgttaattg actacaggaa 240
gccttctgca ttgaaatttg aggtgggtcca tatgccacag gacagaacag aagaaaatga 300
aatattttgtt gacctagctc tgaatgtctt gccaggctta tcaacctggc aatcagttat 360
aaaattaaat gatttttttg ttgaaataag aggaacttta aaaatgatgt gtgagagctt 420
tatctacaat cagacgctta tgaagaagct acaggaaacc aactacgatg taatgcttat 480
agaccctgtg altccctgtg gagacctgat ggctgagttg cttgcagtc cttttgtgct 540
cacacttaga atttctgtag gaggcaatat ggagcgaagc tgtgggaaac ttccagctcc 600
actttcctat gtacctgtgc ctatgacagg actaacagac agaatgacct ttctggaaag 660
agtaaaaaat tcaatgcttt cagttttgtt ccacttctgg attcaggatt acgactatca 720
tttttgggaa gagttttata gtaaggcatt aggaaggccc actacattat gtgagactgt 780
gggaaaagct gagatatggc taatacgaac atattgggat tttgaatttc ctcaaccata 840
ccaacctaac tttgagtttg ttggaggatt gcactgtaaa cctgccaaag ctttgccata 900
ggaaatggaa aattttgtcc agagtccagg ggaagatggt attgtggtgt tttctctggg 960
gtcactgttt caaaatgtta cagaagaaaa ggctaataatc attgcttcag cccttgccca 1020
gatcccacag aaggtgttat ggaggtacaa aggaaaaaaa ccattccacat taggagccaa 1080
tactcggctg tatgattgga taccacagaa tgatcttctt ggtcatcca aaaccaaaagc 1140
ttttatcact catggtggaa tgaatgggat ctatgaagct atttaccatg ggtccctat 1200
ggtgggagtt cccatatttg gtgatcagct tgataacata gctcacatga aggccaaagg 1260
agcagctgta gaaataaaact tcaaaactat gacaagcgaa gatttactga gggctttgag 1320
aacagtcatt accgattcct cttataaaga gaatgctatg agattatcaa gaattcacca 1380
tgatcaacct gtaaaagccc tagatcgagc agtcttctgg atcgagtttg tcatgcgcca 1440
caaaggagcc aagcacctgc gatcagctgc ccatgacctc acctggttcc agcactactc 1500
tatagatgtg attgggttcc tgctgacctg tgtggcaact gctatattct tgttcacaaa 1560
atgtttttta ttttctgtgc aaaaatttaa taaaactaga aagatagaaa agaggggaata 1620
gatctttcca aattcaagaa agacctgatg gggtaatcct gtttaattcca gccacataga 1680
atgttggtgaa aaccttgcta ttttcatatt atctattctg ttattttatc ttagctatat 1740
agcctagaat tccatgatca tgaggttgtg agtatatctc attctttcgt tgcattttcc 1800
taggtgtgct tactctcttc tctcactttg tgacacaagg acatgaatac atctaaattt 1860
tcctatttct gatatcactg tttccatgac gtcattactt ctctaacctt aagtgatagg 1920
gtgacctgca atatgctgat tcctgggtgt tgacacaaaca catggatgta aagaagtaaa 1980
aaatgtaaaa ttcacaaaat tcagtaaacc acacaaatca atgaagcatt ctatgacatt 2040
agcttggttat gagtaacata atgatttttc tttttcaatt taaataagcc cttctacata 2100

```

Sequence Listing

```

cccagcatta ctgatctcag acaatgaatt gctaaaaatg acgatagggc attacactca 2160
gaatagtttg ctatatttcc acatacctca tctagatgtc atagcctaca tttctgccat 2220
cacttaactg acattttttg tgtgttcttg atgataaata gacagttctt attattgtcc 2280
tcaaataata aaagaaactg aaattttctt acatagagaa aatgtccata agatattcaa 2340
gttaaacaga ttatttttgag ataagtaacc attagaaata tgtgattgta atttctgatt 2400
ttataaaatt ttaattgata gtacacttga tttaaatgtc tattctttaa aatgatgaat 2460
actcataatt cttatctcta taatcaaaag tataatttac ttagaaaaaa taaagagatg 2520
cttgttctga aagtaaaaaa aaaaaaaaaa aaaacactgt catgccgtta cgtagcgtat 2580
cgttgacagc ccactgtcat gccgttacgt agcatatcgt tgacagcgac actgtcatgc 2640
cgttacgtag cgtatcgttg acagcactgt catgcgttac gagcgtatcg ttgacagcac 2700
tgtcatgccg ttacgtagcg tatcgttgac agcaaaacac tgtcagccgt tacgtagcg 2759

```

<210> 2
 <211> 527
 <212> PRT
 <213> Human

<400> 2

Met	Arg	Ser	Asp	Lys	Ser	Ala	Leu	Val	Phe	Leu	Leu	Leu	Gln	Leu	Phe
1				5					10					15	
Cys	Val	Gly	Cys	Gly	Phe	Cys	Gly	Lys	Val	Leu	Val	Trp	Pro	Cys	Asp
		20						25					30		
Met	Ser	His	Trp	Leu	Asn	Val	Lys	Val	Ile	Leu	Glu	Glu	Leu	Ile	Val
		35					40					45			
Arg	Gly	His	Glu	Val	Thr	Val	Leu	Thr	His	Ser	Lys	Pro	Ser	Leu	Ile
	50					55					60				
Asp	Tyr	Arg	Lys	Pro	Ser	Ala	Leu	Lys	Phe	Glu	Val	Val	His	Met	Pro
	65				70					75					80
Gln	Asp	Arg	Thr	Glu	Glu	Asn	Glu	Ile	Phe	Val	Asp	Leu	Ala	Leu	Asn
				85					90					95	
Val	Leu	Pro	Gly	Leu	Ser	Thr	Trp	Gln	Ser	Val	Ile	Lys	Leu	Asn	Asp
			100					105					110		
Phe	Phe	Val	Glu	Ile	Arg	Gly	Thr	Leu	Lys	Met	Met	Cys	Glu	Ser	Phe
		115					120					125			
Ile	Tyr	Asn	Gln	Thr	Leu	Met	Lys	Lys	Leu	Gln	Glu	Thr	Asn	Tyr	Asp
	130					135					140				
Val	Met	Leu	Ile	Asp	Pro	Val	Ile	Pro	Cys	Gly	Asp	Leu	Met	Ala	Glu
	145				150					155					160
Leu	Leu	Ala	Val	Pro	Phe	Val	Leu	Thr	Leu	Arg	Ile	Ser	Val	Gly	Gly
			165					170						175	
Asn	Met	Glu	Arg	Ser	Cys	Gly	Lys	Leu	Pro	Ala	Pro	Leu	Ser	Tyr	Val
		180						185					190		
Pro	Val	Pro	Met	Thr	Gly	Leu	Thr	Asp	Arg	Met	Thr	Phe	Leu	Glu	Arg
		195					200					205			
Val	Lys	Asn	Ser	Met	Leu	Ser	Val	Leu	Phe	His	Phe	Trp	Ile	Gln	Asp
	210					215					220				
Tyr	Asp	Tyr	His	Phe	Trp	Glu	Glu	Phe	Tyr	Ser	Lys	Ala	Leu	Gly	Arg
	225				230					235					240
Pro	Thr	Thr	Leu	Cys	Glu	Thr	Val	Gly	Lys	Ala	Glu	Ile	Trp	Leu	Ile
			245					250						255	
Arg	Thr	Tyr	Trp	Asp	Phe	Glu	Phe	Pro	Gln	Pro	Tyr	Gln	Pro	Asn	Phe
		260						265					270		
Glu	Phe	Val	Gly	Gly	Leu	His	Cys	Lys	Pro	Ala	Lys	Ala	Leu	Pro	Lys
	275						280					285			
Glu	Met	Glu	Asn	Phe	Val	Gln	Ser	Ser	Gly	Glu	Asp	Gly	Ile	Val	Val
	290					295					300				
Phe	Ser	Leu	Gly	Ser	Leu	Phe	Gln	Asn	Val	Thr	Glu	Glu	Lys	Ala	Asn

aatgtttag	ctgtttttca	taaagaaagt	gtatttggtc	taagcgtaag	ataacctact	1080
ttcttaatac	cagtaataata	cttaaaaaatg	atcatcaata	actaagagat	tataattttgt	1140
atttcoctcca	aatagcgcaa	atcaacatca	catatttttg	agaatcactg	attgttagtc	1200
tgaatgttat	agaattttcta	ttgaaataaaa	atgctaataca	ttattttctc	tctcatcatg	1260
tatttaagaa	aatcttcaga	aggctcttctt	tgaattaatt	tttcaagagt	cattaaattg	1320
aacattttct	agaatttctt	aatttcttag	gtgattactt	cacaaaaact	tgaaaaaata	1380
ttataaaaaag	ttaaaaaact	tacggctcttg	tggggcataa	gatagtagaa	tttttacttt	1440
actgatatac	acctatttga	cttatttttta	tttctttgct	ttactgataa	aaagtgtgtt	1500
tgcctttgcaa	ttttcatata	gttgatgata	gagctgggtca	atgcaagaca	tgtttttatc	1560
caaatatgtt	tgagaattat	gtagaacat	gaaaaaagg	acaattatat	cgcacactaa	1620
aatattgttt	aatgtattcc	aacgaattct	tatgcataga	ctgtttcaca	gaactaatat	1680
tcagaggatc	ccagttcaaa	tgtccttagc	cttagacatg	atttgaattt	acatatattg	1740
atttgcttta	aataattttc	cattcagtaa	gctgtgccta	gctgcagata	gcctaccagg	1800
ctttatggat	ctaggttaaac	aatacaaatc	tcttggtcctc	aagtctacat	tcagatatta	1860
atttaaagg	gtacagctat	atagagggtca	ctggcaaatt	ttggtaaaat	aggattatag	1920
taaaagcccc	ctgacaagat	tgaaatttaa	aataaaacaa	aagtgttatc	aaaggggtga	1980
aagagcattt	tccaataaac	aaaagtgggt	tctggccatg	cattcagaaa	ttccccaaca	2040
attcttttaa	aatcatggag	cagcttgata	tataagaaat	tcatttaata	actatattta	2100
ttatgtagct	ccaacttact	aaattattga	ttattatata	ttttatagaa	ttatctattg	2160
tgagtctaaa	tcaagagtat	atattcaaac	aactatagga	aaagggatat	cagtcaattt	2220
caattcaagg	atattatttc	ataagtgtt	acgcacaggt	gtatttcatt	ttattatata	2280
ttgctttatt	gtccttcaca	aaaattgcaa	tttacaatt	aaaggttttt	gaaaaccttg	2340
aatcaagcta	atcaatttgg	cgtaatat	ccaacaacaa	gtgtgtactt	ttgactctat	2400
cacatattgg	catttatcat	gctttttcaa	atttttcatt	gttatatctg	ttacggtgat	2460
ctgggatcag	tgttccttga	tggttacacg	tttattagct	tgggggcacc	ttgatgtgtt	2520
acaatataag	acagcaaaact	taattataaa	tgttgtgcat	gtactaactg	ctccgctgat	2580
tcgtttcccc	atccccactt	ttcttaggcc	tccctattcc	ctgagacaca	gtaataaac	2640
atacaatgac	ttctaaatgt	tccagtga	agaaaagtag	caggtctctc	aatttaaacc	2700
aaaaatataa	aggaataagt	ttaatgagta	ctatagttaa	gatatggtt	gcttgaccct	2760
acaaaatcct	gtgttgaaat	ttgatcacca	atattggagg	tggggcctga	tgggaagtgt	2820
tagggctcatg	agggtagatt	ccttatgagt	acattaatgc	tctccctggg	gaaatgggtg	2880
agttcggtt	cactctatta	ggtcccagga	gagataatta	ttaaaaagag	ccaggaacat	2940
ccaccttctt	tctcttgcat	atctctcatt	atctgatccc	tgcacttgct	ggctcccaac	3000
atcttcttca	atgagtggag	gaaaccagag	gtcttcacca	gacacagatg	ttggtgccat	3060
gcctcttgta	tacctgaag	aattgtgagc	caaataaaaa	cctttttctt	ttacaaatta	3120
gacagcctca	gttattcctt	tgtagcaaca	aaaaaagcct	gggacaggcc	aaaaactaca	3180
ccattgcaac	aaacagttta	acaagatgtg	agtgcacaa	aaaagttttt	ggaggaaatt	3240
aaaagtgtca	ctccagtgtg	catacaaatg	ataagaacaa	ataaccatta	tcagtgtctga	3300
tatggagaaa	attttagttg	tctggagaga	aaatcaaatt	agctagccag	ctgcagtgat	3360
tcatatctgt	aatcccagta	acttgggagg	ctcaggtggg	agaacggctt	gagcccagaa	3420
gtttgaagtc	caaggctgca	gtgagctatg	attgctccac	tgcactccaa	cctaggtgat	3480
agagcaaaac	cactaccaaa	aaaaaaaaaa	aaaaaaagaa	gaaaaagaaa	agaaaaaaa	3540
ttaaaccaac	cacaacatca	ccttaggttt	tggcattagc	taaaaactaa	tacatagtaa	3600
agcgttaact	attcaattgc	atgaaaacct	agagaggaga	ggaagatgca	gaaaaaaaga	3660
ctgaagctag	tagaggttga	ctaattgaggt	ttacaggaat	aaactgccta	catgatgcaa	3720
aagttcaatg	tgaagcaata	ggaagtcatg	cagaagactt	agctaataata	ctcagtaaat	3780
gtggctacag	taaacaaatg	attttcaatg	tagacctaac	agccttctgt	tggagaaga	3840
tgccatttaa	aactttcata	gctagagaag	agaagtcaat	gcttgtctct	gaagctacaa	3900
aaaacaggct	gaatctcttg	tagtggctaa	tgcagctgat	gacaaaggta	aagccaatgc	3960
ccatttactt	tttgtaataa	ttatagagga	ctcttaataa	ttatgttaaa	tctactttgc	4020
ctgtgttata	tcaatggaac	aacaaagcct	ggatgatata	acattgggtat	atgacatggc	4080
ttattgaata	ttttaagcac	actgttgaga	cctattgctc	aaaaaagagg	attcctttca	4140
aaatattgct	gctcattgac	aattcacatg	gtcaacaaag	ggctctgatt	aagatgtaca	4200
gatattaatg	tttgccctgct	tgctattatt	acatccatct	tacatgccat	ggatcatata	4260
gccttgactt	tcaagtctta	tgtaaagaaat	atattttgtg	aggctatagc	tcttactaat	4320
ggggaaagta	tattgaaaac	cttttcaaaa	ggatttttca	ttctagattc	cattaagaac	4380
attcatgggt	catgagagga	agtcaagata	ttaacattaa	caagagtttg	gaaaaaattt	4440

gattctaact	ctcctggatg	atthttgaggg	attgaagaca	tcatgtgaag	aattaactgg	4500
ggatgggggtg	gtcatgaaaa	aataaataga	attataagtg	ggcctgaagg	tttgtctaaa	4560
ttgtctataat	atcatgataa	aactaaaacc	tgtaaaaccg	gtgaggaggt	gcttttttaa	4620
cagttactttt	ttatagatga	acacagaaat	tggttttgtg	agttggaatc	ttctccgagt	4680
gaaaatgcta	tgaacattgt	tgaaatggct	acaaatgact	tagaatatta	cacaaaatta	4740
gtagataagg	cagcatcaag	gtttgagaga	atggactcaa	atthttgaaag	aaattctact	4800
atgggtaaac	tgctgtgaaa	catcatcata	tgctacagag	aaatctttca	tgaaaagatg	4860
agtcaattca	tgcaacaatc	tttgtttgtct	aattttaaaa	attgtccagc	tgccctgatc	4920
aatcaacagct	aatcagcact	gaggcaagac	cctacaccag	aaaaaataaa	aataaaaaaac	4980
ctcacttgct	gaagactcag	cttattatta	gcacttttta	gccatacttt	taactaaggt	5040
atgtgcattc	cttttttaaac	gtgatgatat	tgacacagcta	atagcctaca	aggtatgggt	5100
aacataactt	ttatatgtcc	tgggacccaa	atthttgtgtga	atcacttttat	tgacatattc	5160
ctttttattga	gatgaactgc	aacttatctt	gcaatatctc	caagatatgt	gtgtatggca	5220
tttcaaataa	gatgtgaaat	tattttatta	gtataaaaag	caaatttaat	tttctttcct	5280
ttgatcatct	ttatccttgt	tactgtgtat	ttatccttta	aacattgaat	gactccaatt	5340
gtttaaaact	gagtctttct	taaatgagtc	ctaatatcat	agtaattaaa	atcacctaca	5400
agttggtaat	gcaggcagca	tgtgaggcac	agaaaacaac	aaattttataa	gacataaatg	5460
catttgcttg	gaagctgaga	gaaggctcta	ttctaatttc	tgataacttc	aaactgagta	5520
tcttcagtaa	aattttattca	ctatcaaatt	caaggcgctt	ggattttatga	cctaggaaaa	5580
aacttcaaac	attaaaatgt	gatgacctta	aaaagaggct	ctccacacta	tggtgtataa	5640
caccaccaac	tttgattaga	atthttaaaga	gaaacaaatt	ctcttatgga	gtttatcttt	5700
ttatcacttg	caaaatatgt	ttttgtaaaag	agataactaat	tacttagtta	tttgtagtta	5760
gccattcttc	tgattaaaaa	cctaaaatta	aatcttgaaa	atgtgttttc	cttcaaaaac	5820
catcatttga	gagaaacact	aaagtaagtg	tatgattatc	atagcatgta	cataggtgct	5880
tcacaaccca	aaaagaatat	tgtcatgggt	aagaatcagt	aaaggaattt	ctcctaataa	5940
aacagtagcc	tattgaattaa	agtaattgata	tgcataacag	caagttaaag	ggaactgatc	6000
ctggtgggac	tattgaaaga	tatacccttg	actatagatt	agaaaataca	gagatgttat	6060
ttagtgaaga	tattgtggta	ctcattttatc	atctgcaatt	cacttgcaag	ggaaaaaatg	6120
agtaataaat	tcattttgcat	tttggaattg	tgtcttttaag	ttgtgaaaat	acacttaaat	6180
ataaccatct	gtccttttgt	ccttccttcc	ttccttcttt	ccttccttcc	ttccttcttt	6240
ccatccttcc	ctccctccat	ccttccttcc	ttcctttctt	ccttccttcc	tttctttctc	6300
tgtccttctc	tcttttttcc	tttctttctt	ttttcttctt	tattatttca	tttaattccc	6360
cttccatttg	acgtctaaaa	gccatgttgt	tctagaggac	ttaaaacttat	ttttttctta	6420
atagcttact	gaaaaattag	tgatacaatt	ttttatttga	attgtatgct	aattcattct	6480
gttattttct	ttattgagga	aggcccacta	cattatgtga	gactgtggga	aaagctgaga	6540
tatggctaatt	acgaacatat	tgggattttg	aatttcctca	accataccaa	cctaactttg	6600
agtttggttg	aggattgcac	tgtaaacctg	ccaaagcttt	gcctaaggta	ggactattgt	6660
attaaggaat	attatgtact	ttatgacatg	acttgttttc	ccttgaaaaga	ttacaacctt	6720
agttatagaa	ggatgatgtt	gaatgtcgtc	tgthttgcagc	tccatattta	ttttccatgc	6780
cacaggggct	cttataggtg	attatatgtc	ttttcggtat	tatattgaga	aagtaggcag	6840
aagaatttca	tgattagaat	agatttttaa	atactagtat	tacaatagtt	tggaataata	6900
attgaattaa	tagggaattg	gagccatgaa	gatcactaaa	aagaatgctc	tagcctttct	6960
cacaatcaaa	ttgggcttat	gaacaaggat	atthttgtcatg	atagtacaga	aataagcata	7020
ttttcatgag	acatatgtga	tatatccac	aggagttgggt	gagtgaagaga	aaataagtga	7080
tgaaggaaga	caaagaataa	aagaaaattt	caataaatgg	aaagtthtaag	tgthttaatga	7140
tagtgatgac	ttttactcaa	ataagtgtct	agaagtcac	ttgtttgtga	tttatatgat	7200
gaattctgtg	ttgtgactat	ccactttgag	ctcgtgagaa	tgthtaggtga	ggthtaataa	7260
aagccatttg	agaaaaacaa	ggthttcaacc	tctgtggaca	gaaatctaaa	tatcgatagt	7320
tatcaggaca	aagtagagct	catagaaata	atthttgcagc	ctgcaggtht	gthttggagt	7380
gaaaataaaa	ttgtatacta	tattcctaaa	tcatcagagg	aaaaaattta	tagthtcaagg	7440
aatgttgaaa	gaaacaatat	tgagaagtaa	aagtgaagtaa	tagthtgtht	agthttttta	7500
tagthtttgta	agtatgtctt	gagthtactg	tcccaaaagt	ggctatttag	tctagccttg	7560
acctgacaag	gttctaggat	atthtagtcat	ggatgttcat	aatctacctc	ttacgggata	7620
ctttttattc	tgatgaacag	cctaattgct	aagtgtgcaa	tctataccaa	gattgttctt	7680
atagggaaact	tgthttacact	ggaagacacc	actgtgtctc	ttgtatgacc	tatgtcttct	7740
ttatccctac	aaagthaacc	acatttatagg	aaacctgtac	aaggccagat	gttatatttg	7800
tgthtggtcaa	gtgagaaaaac	atggggagaaa	cttaacccaa	cacataaaaat	aacagaaaca	7860

gtcttctttg	accattttcta	gagaaaagag	ttcagcatcc	cttgtaaggc	cactaggaag	7920
aagaaaattc	tctgggaaaa	gcacattcaa	ccaatgaatg	gagaccaaga	aagagagtga	7980
gggatctatg	tgccaaaatg	ttaactggga	tccagggtgt	tacctagggtg	ggtttccaat	8040
ggggaactgt	aattggtagg	tttaatgcaa	gcaggcacaa	agtcocatgga	ggcatttctga	8100
gactgaaaga	tagtcacttt	ggcatatctg	cacagaatct	gatcagtgat	tcaagcccaa	8160
gtaggctgta	tctagttgtc	ctatagggtg	gttaccagga	ggcagtggtg	aagtaaaaaat	8220
cctgactgaa	cacattgagg	aaatggaagg	aggtggaaga	ttttaaacgg	tgtcagtggt	8280
gactaagacc	tgcttctggg	atggaaaatt	caacttatat	tttaaatgca	tagccagaca	8340
acataaaaatt	ataagaatth	accacaatag	ctatggtaac	aatactgggt	ttacctatta	8400
ctacagagtg	aaaagaaaaac	cctcattttcc	cattttatgg	aaatataatc	aaaatcctat	8460
aaggaagggt	tcagagccag	taggattttcc	agaaaaatta	ttggtttttat	agtaagatgt	8520
gtattgatga	atataattht	atttattaat	tattaatatc	actttactta	ccaggaaagt	8580
tataccagaa	aaccaagctc	tcttaagcca	tggcatctgt	atctaaaata	gaaatacaga	8640
aggagagctg	acaattttcca	tcatttctcta	ggtaatctcc	catgccattc	taccctttat	8700
tcccacactc	ccagtttttac	acacacacac	aaacacacac	acacaaacac	acactcatag	8760
aaataatcat	agaagacata	tttttaaaaa	agttagatcc	atacagtaat	aattttattag	8820
gtaaaagctt	ttgtgctgat	aattttacaa	gtttaattga	gataattht	agggctgtct	8880
tacactaaat	atthatttht	atthtttaaa	tttgacatgt	aataattgca	catgtthtaag	8940
agaaatgctg	tggtattaca	atacattthaa	atgttggtga	ataattacat	caagataata	9000
aacccatcat	ctaaatatth	atcattttctt	tgtggtgata	acattcaaaa	acctcctthc	9060
tggctatctt	gaaatatgta	atacattact	attaactata	gttaccaca	aacttaatat	9120
aataacagaa	catatttctt	caaattthaaa	cgttgtatcc	attgatccac	cattttctcat	9180
tgccctccct	actatctctt	cagcctctag	taaccacaat	tctactctct	aattatatta	9240
tgaatgcatt	ttttgattcc	acatataagg	gataccatgc	tatctctgcc	tggattatth	9300
cagttaacat	tatgccctgg	aggttcattc	atgtttctac	aatgacagg	atttcattct	9360
tttttttcca	atatatatth	aatgaaatgg	atatataata	acattggaaa	atgtatatat	9420
atatatatat	ctccagtggg	atgctattga	gctataaaaa	agttaatata	taatagaaat	9480
aaagcttata	tatatctaata	ggaatggata	tatatatata	atggaataga	aatatatatc	9540
tatacatata	aacacacgca	atatacatat	ccattttcatt	gcataatata	atatatagag	9600
agagagagag	agagatatth	tcaaagtgtg	gtatatatat	ccaatggaat	ggacatatat	9660
atatgtatat	tttttccata	ttttcttttat	gtattttctt	attaatggat	gtttagggtg	9720
attcatccct	tgggtatatg	aataatgthg	atgtaaacad	agaaggacag	atatctctat	9780
gacttcttag	tttatthaaa	tatacaccca	gtaatggaaa	tgctgtataa	tatggtagtt	9840
ctattttcat	tttttgagga	actaccatac	cgthtttctt	actaattgta	ctaatttgca	9900
tttccctcaa	cagthttataa	aagatcttct	ttctctgcat	actttctagc	acttgthatt	9960
tttgccthtt	gataatagcc	ataacagggg	tgatgtgata	tctcattgta	gttttgattt	10020
gcattttccct	gatgattagt	gattttgagc	atthttgtaat	tatacttctt	agtcactgat	10080
agtcttctth	tgagaagtgt	ctattcaggt	ctthttgctta	ttthtttaatc	aaattagtaa	10140
tttatthttta	ttgactgatg	tgacttctat	gtatatthtga	gatagtaact	tattgtcaga	10200
ttcatagtht	gcaaatatth	ttcatgttgt	gaattgtctc	ttcacctgt	tgthttgcttc	10260
atthttctctg	cacaagctca	atgctthtgat	ataacccatt	tatctactth	tcctthttgth	10320
ggctgtgctt	ctgaagtccct	atccaaaaaa	atccttgccct	agaccaatgt	cacaaatcat	10380
tcctcttaca	gtthcttcta	gtagthtgat	aatgthttggc	cttatatthta	actthtgtaat	10440
tcattthttac	ttactthtga	tatgggtgagg	gatagaggtc	tagthttcatt	ttctgcatgt	10500
ggatatgcag	ttttccttag	accatttagt	gaagaggthg	cctthttttct	attatgtgtt	10560
cttggcacct	ttgtcaaaaag	tcagthtagct	gctatatthcc	tccattthtg	ttgttataga	10620
ggaacacatg	agactagcaa	atthtatatat	caaatagaaat	tattthgaatg	atagthtctgc	10680
atactgtaca	agaagcacag	cactgactth	tgctthggcct	ctggtaagggt	tctcaagatg	10740
cttccactthg	tggtagaagg	caaacatgag	ctgggtatatg	caaaggthctc	atgacaagag	10800
aggaaaccat	aaagagggga	tgtgaggggag	tgccaggtht	tgtaaaaacaa	ctagctcttc	10860
tggaactaa	tagagtaaaa	atthgcctcc	caggcagggg	attaatctat	tcatgagggg	10920
tctgctthcca	tgacaaaagg	acattctgtt	agattctacc	cccaatattg	gggatcaaat	10980
thtaacatga	agtgtggagg	gctcaaatat	ccatactatg	gcagcagtaa	atgcataaat	11040
thattthtthg	gatctctatt	ctatatagta	thgggtgatg	tatctgthtt	catgccactg	11100
ccatactgth	thgggtgatga	tatctatgct	atatatgtgt	gtgtgtatat	atatattata	11160
tatatgtata	tatgtgtata	thatatatat	gtatatatgt	gtatatattata	tatatataat	11220
actthtaagth	thatatatat	ataaaaact	thtaagthcaa	gggtacatgt	gcaggatgtg	11280

caggtcagtt	acataggtat	acatgtgcca	ttttggtttg	ctgcatgcat	caactcatca	11340
ttacattagg	tattttctct	aatgctatcc	ctccaccagc	cacccaaccc	ccaacaggcc	11400
agggtgtgta	tgttccocgc	cattgtgtcca	tgtgtttctca	ttgtttcactt	cctacctaaa	11460
agtgagaaca	tgcaagtgttt	gattttctat	cottgtgata	gtttgctgag	aatgactggt	11520
ttcagcttca	tccatgtccc	tcaaaaggac	atgaactcat	cottattttat	ggctgcatag	11580
tattccatgg	tgtatatgtg	ctacgttttc	ttaatccagt	ctatcactgt	tggacatttg	11640
ggttggttcc	aagtctttgc	tattgtgaat	agtgtacaa	taaccatatg	tgtgcatgtg	11700
tctttatagc	aacatgattt	actatccttt	gtgtacatac	ccagtaatgg	gataactggg	11760
tcaaattggt	tttctagttc	tagatccttg	aggaatcccc	acactgtott	ccacaatggt	11820
tgaactaatt	tacattccca	ccaacagtg	aaaaacgttc	ctatttcccc	acatcctctc	11880
cagtatctgt	tgtttcctga	ctttttaatg	atggccattc	taactcacat	gagatgggtat	11940
ctcattgtgg	tttttgtttg	catttctctg	atgaccagtg	atgatgagca	ttttttcatg	12000
tgtcttttgg	ctgcataaat	gtcttctttt	gacaagtgtc	tgttcatatc	ctttgcccac	12060
ttttcaatgg	agttgtttgt	ttttttcctg	taaatttggt	taagttcatt	gtagattctg	12120
gatattagcc	ctttgtcaga	tgggtagatt	gcaaaaaatt	tctcccattc	tgtagggttg	12180
ctgttcaccc	tgatggtagt	ttcttttgct	gtgcagaagc	tcttttagctt	aattagatcc	12240
catttgtcaa	tttcggcttt	tgttgccatt	gcttttggtg	tttttagtcat	gaaacccctg	12300
cccaggccta	agtctcag	ggtatagcct	aggttttctt	ctaggatttt	tatggtttca	12360
ggtctaacat	ttaagtcctt	aatccatcct	aaattaattt	tgtataaga	tgtagaagg	12420
gatccgtttc	aactttctac	atagtggtag	cgtgttttcc	caacaccatt	tattaaatag	12480
ggaatccttt	ctccatttct	tgattttgtc	atatttgtca	aacatcacat	ggttagagat	12540
gtgatgtgtt	atcactgagg	cctcttttct	gactccattg	atctatatat	ctgttttgat	12600
accaatacca	tgttgttttc	gttactgcaa	ccttgtaatg	caatttgaca	ttcaggacca	12660
tgatgcctcc	agtctctttt	tttttttcta	aataattttt	ttgtcaatgt	aagctcattt	12720
tcgcttcttt	ctgatccata	aagtattttt	ttcccatttt	gtggagaacg	ccgcnnnnnn	12780
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnggcaca	12840
cctcgtgcgc	atatatatat	atatatatat	atatacctct	atatatatat	acatacatat	12900
atacatacac	acctccttgt	ctgggtgtggg	atcagggtaa	tgctagcctc	acaagatgat	12960
actgaagtgt	ttttgccttt	ttgacttttt	gatggtttgg	aagagtgaga	aaaagtgtta	13020
tttaattatc	tttaaatttt	gttgaatttc	atagtgaaga	ccttagctca	ctggcctttt	13080
taatgagaac	tttattactg	atttaaactt	cttcttcatt	atttatttct	gccttgtttt	13140
tattttcttca	taatccagtc	ctatttttatg	tgtccactaa	attgtttatt	ttcctagaat	13200
ttttccattt	attggcatat	gcatgtccat	agaagccttt	tatagtctct	ttcattttcta	13260
gtgtcatttt	tttccttttt	tttaagaatc	cttaagattt	tagagatgaa	atgtcacttt	13320
gttacgcata	ctggagtgcg	gtgacattat	tatagctcac	tgaaccctaa	actcctgagt	13380
ttaagcaatc	cttctacctc	aaaattccaa	aattcctgag	tacgtgagac	aggcatacac	13440
catcaagact	ggctaattta	tttcaaattt	gttagagatg	ggttcttact	aagctattct	13500
caatctttgg	gcttcaagtg	attcttcagc	ctctgtctct	gaaaatgctg	ggtttataga	13560
tatgagcctc	tatgocctgat	ttgctttgtc	tctttgtaat	ctcccatttt	atttgtgtct	13620
ttctcgtgtt	gtttcatttt	gttatgtttt	cagttacctt	gctaaagctt	tgtcgatttt	13680
atctcttcaa	acaactaaat	caatatattt	ctgattttcc	atatagtatt	ttatttctat	13740
ttcattttatt	tctgctctaa	tctttgttaa	atatcttgtt	ttcctaataa	ttttgagttt	13800
ccttgttctt	gttttctaatt	tccttgogat	gttatcataa	attgtttatt	tgatatcttt	13860
ctactttttt	gatgtgtgtg	ttcgttggtg	tagactttcc	tctttattat	tctgatttct	13920
tcctcaattc	tctaataatta	tgattgcatt	attttccaag	tttcttttgt	ttttttattt	13980
atagtttatg	tgattcctga	acttgtcaaa	gagattattg	tgaatttgat	gtcggatatt	14040
taagcatttt	caaaaactttt	ggtgcattat	tgaattttta	ttggtttatt	ttagagatgt	14100
catacttccc	agtttttttt	taacaatact	tgtcttttat	attgatgtct	acataatttaa	14160
aaagataacc	acctgattca	gcttttttaag	gtgatatgca	gtgggtgttaa	gtgtgtactg	14220
cttaatatca	gagctgaatc	actgccctga	ggattctttc	tgttctgagg	agagcttgta	14280
gttaatagca	gaacctaaat	agtgcagtag	agctaaatct	cttccatgct	gttggttttc	14340
tgtctgggga	agacttatca	tgaccatgaa	aacataatgc	tgtgccagaa	cttaaaccca	14400
aacctgtagt	aattttctgag	ttgaggaagg	cttaagaaat	aactggaact	tagttactaa	14460
cctgatagtt	gtttctgagt	cagagaaatg	ctctgcatga	tcacctggga	tatttgtaaa	14520
atctaaccaa	agattctagc	cttccccttg	gatttgtgct	cctgtactac	tgtagtgtctg	14580
gctaggtcct	catcagtgaa	ttccctgctg	ataggaccac	aaagcatctg	ccaagatctg	14640
tttgccattt	gctgtgatta	gtgcttctgc	tctttgcttc	caattcaact	caggtgggtc	1

agcccttctg	acactcctaa	tacctcctgt	gggatggaac	atagaaggct	tctcacaatg	14760
attcacacac	tgatatggag	attgaaatgtc	cagttgcaac	tattttcttc	cacctgtgta	14820
attgcaggta	cagggaaagt	ttctgtgact	gatgctattt	tggtttgagg	aatgggggtga	14880
tgtggcacaa	tgtactttct	tctttctggg	catggatttt	ttaattttcca	tgaacccata	14940
agatttttca	ctttttctct	gagctctggg	gctttcagag	tggatatttt	atattcgaat	15000
ctgtgctagt	tgtactttta	aaagcggattg	atgctggagg	tcttctattc	caccatctcg	15060
ctgatgtcag	tctcaaaata	ataattttat	attttagcaa	attattttgg	ttttaggatt	15120
ttgtgtctac	gtgacacaga	catgaaaaga	gatgtactca	ttactgaaac	tttttgcata	15180
ctgttttggg	tgtgcgcctt	ttctagtagt	aatgattaca	tatttaagcc	acatgtttta	15240
tacatagact	gtccttttaa	gagactagat	agttctgtgt	gtcagcatat	agggacagaa	15300
tataactaca	cattaataat	ttctcaagta	tttatttttag	aagtgtaaat	aacctttatt	15360
ttaatTTTTG	ttatattatg	cctctgtaat	gcagataaat	ttttatcttc	aggaaatgga	15420
aaatTTTTgtc	cagagttcag	gggaagatgg	tattgtgggtg	ttttctctgg	ggtcactggt	15480
tcaaaatggt	acagaagaaa	aggctaatat	cattgcttca	gcccttgccc	agatcccaca	15540
gaagggtcagt	aaaacctcca	atcctgataa	gcagctattc	acataatgaa	acagtatggt	15600
tttatttggg	tcttgaatct	cattttccac	ttagcataac	aggtaccaa	atttgcaaaa	15660
cattatagta	gtgtacatgg	gcataactga	tcatttgcc	actgagtctt	gctgttactg	15720
gaaacaactt	tcttgattgt	catttgttta	taataaaaata	gatataataa	ataaagctct	15780
accttatatt	ttaggatttg	aaatctaaaa	gcgtgtgcc	atgattccaa	aaaaaaattc	15840
tgacatctat	tatttcaaag	gaccagaaaa	aggaaaactg	atataaaaaa	aaaaagaaga	15900
atcaatctca	agaatatctt	ctcatatttg	tgtgtataaa	aactgtattc	agggtagttt	15960
tgcttagaaa	taaaagctca	gattaatgta	gtctttctaa	ataattagaa	gtttcaaaaag	16020
taaaatgtca	attacaatta	tagtatagta	acaattattt	aagtaatgta	attattttatg	16080
atactccact	aatttttaact	ttattattac	tgtaatctca	gaatttcaca	ctttagatag	16140
tgctatatat	aaactatcca	aaagatatatt	cattttatat	ttagctaaaa	tacttcaaac	16200
tcaataaagg	caagcatact	aattaggaat	ttgaaatatt	gtaatttcaa	ttatgaaatt	16260
atctgttaag	tagtttgaaa	catctatgcc	gttctttgtt	ttcaaagtga	taaaatttgt	16320
atagggtgtcc	aacaaagaaa	aattgtgtaa	aaaaaaggta	caatctcaaa	gaaaatttat	16380
cattgaacag	tggaaacata	gtaattttct	agctcattct	tcttcaataa	aacaattaaa	16440
tataagaaga	aagaggccag	gaaggaaaata	gagaagaaaa	gacacccgat	tatccaaaag	16500
acacacataa	ttgaaagcaa	atTTTTatct	gcagggaact	gtaaatttga	tggtagaatg	16560
agattggctc	catgagttaa	aatgacacac	agatcaggta	cttataaaat	ttttaattct	16620
tatataaaaa	tagattagcc	actgctgaat	tattttttta	aatattcact	ggtattctca	16680
ttctcaaaata	tttttaattg	gtaataaaat	aataatagca	tacctaatag	gcaactggta	16740
cacattattt	taaaagatct	ttgtaaaacg	tctactata	tctttcagtc	tttacgcggt	16800
agctctacac	acccctgtct	caacctatc	ctgaagtaca	atgagtttat	aatttataac	16860
tatatctaca	tccttagaat	gctaataatc	tgtggttcac	tctgtgaaat	acatgtgttt	16920
cttcogtagg	tgttttaggag	gtcaaaagga	aaaaaaccat	ccacattagg	agccaatact	16980
cggnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	17040
nnnaaaaaaa	aagacccaat	cccaaagaaa	atttatcata	gaacaatgga	acataggtaa	17100
ttctctagct	cattogtctt	caataaaaaca	aataaatata	agaagacaaa	ggtcaggaag	17160
gaaatagaga	agaaaagata	accgattatc	caaaatcaca	cacaaaattg	aaagcaaatt	17220
ttatctgtgg	ggaactgtaa	atttgatggt	agaaccagaa	tagttccatg	atttgaaatg	17280
acacagagat	catgtactta	taaaatatatt	tattcttata	agaaaattga	gtagccagtg	17340
ctgaattact	ttttaatgat	tcactgatat	tctcatactc	agatatttta	attgatatta	17400
aaataataat	agtatactta	atagtcaact	ggtacacatt	atttgaaagg	actttttgtaa	17460
aaagtccctac	tatgtctttt	actgtttaca	cagtacctct	acatacccct	gtctcaacca	17520
acacctgaag	tacaatgagt	ttataattta	taactatatc	tacatcctta	gagtgtcaat	17580
atcctgtggg	tcaatctgtg	aaatacatgt	gtttcttcca	taggtgttat	agagatacaa	17640
tggaaaaaaa	accatccaca	ttaggaacca	atactcggt	gtatgatggg	ataccccaga	17700
atgatcttct	tggtaggtct	atgagaaagt	aaaaatatga	actagacgag	gaaaaaatga	17760
ataaatgtta	aacagcaagc	aaattcagca	aagatctaaa	attataaaac	tttattttac	17820
ttactctttt	gaagcagata	taattaaagg	attgactaaa	attgtataga	ttcacacttt	17880
ctattgttaa	ggtgagagtg	acaggaatt	cagaaggaat	taatgcctat	ttttctggag	17940
atagaaatga	tctttagtag	caatgtcca	tgtgtccacc	ttctaaagaa	agtgtgttac	18000
gcttcagtga	gttatctcgt	aattcccatc	tgtagttttt	aaataatttt	aaaagtttag	18060
aataaaaatat	ctcaccattt	ctcatccaat	ttacatacta	ggtcatccca	aaaccaaagc	18120

1													5													10													15						
Leu	Asn	Val			Leu	Pro	Gly	Leu	Ser	Thr	Trp	Gln	Ser	Val	Ile	Lys	Leu																												
																		20													25													30	
Asn	Asp	Phe	Phe	Val	Glu	Ile	Arg	Gly	Thr	Leu	Lys	Met	Met	Cys	Glu																														
																		35													40													45	
Ser	Phe	Ile	Tyr	Asn	Gln	Thr	Leu	Met	Lys	Lys	Leu	Gln	Glu	Thr	Asn																														
																		50													55													60	
Tyr	Asp	Val	Met	Leu	Ile	Asp	Pro	Val	Ile	Pro	Cys	Gly	Asp	Leu	Met																														
65																			70													75													80
Ala	Glu	Leu	Leu	Ala	Val	Pro	Phe	Val	Leu	Thr	Leu	Arg	Ile	Ser	Val																														
																		85													90													95	
Gly	Gly	Asn	Met	Glu	Arg	Ser	Cys	Gly	Lys	Leu	Pro	Ala	Pro	Leu	Ser																														
																		100													105													110	
Tyr	Val	Pro	Val	Pro	Met	Thr	Gly	Leu	Thr	Asp	Arg	Met	Thr	Phe	Leu																														
																		115													120													125	
Glu	Arg	Val	Lys	Asn	Ser	Met	Leu	Ser	Val	Leu	Phe	His	Phe	Trp	Ile																														
																		130													135													140	
Gln	Asp	Tyr	Asp	Tyr	His	Phe	Trp	Glu	Glu	Phe	Tyr	Ser	Lys	Ala	Leu																														
145																			150													155													160
Gly	Arg	Pro	Thr	Thr	Leu	Cys	Glu	Thr	Val	Gly	Lys	Ala	Glu	Ile	Trp																														
																		165													170													175	
Leu	Ile	Arg	Thr	Tyr	Trp	Asp	Phe	Glu	Phe	Pro	Gln	Pro	Tyr	Gln	Pro																														
																		180													185													190	
Asn	Phe	Glu	Phe	Val	Gly	Gly	Leu	His	Cys	Lys	Pro	Ala	Lys	Ala	Leu																														
																		195													200													205	
Pro	Lys	Glu	Met	Glu	Asn	Phe	Val	Gln	Ser	Ser	Gly	Glu	Asp	Gly	Ile																														
																		210													215													220	
Val	Val	Phe	Ser	Leu	Gly	Ser	Leu	Phe	Gln	Asn	Val	Thr	Glu	Glu	Lys																														
225																			230													235													240
Ala	Asn	Ile	Ile	Ala	Ser	Ala	Leu	Ala	Gln	Ile	Pro	Gln	Lys	Val	Leu																														
																		245													250													255	
Trp	Arg	Tyr	Lys	Gly	Lys	Lys	Pro	Ser	Thr	Leu	Gly	Ala	Asn	Thr	Arg																														
																		260													265													270	
Leu	Tyr	Asp	Trp	Ile	Pro	Gln	Asn	Asp	Leu	Leu	Gly	His	Pro	Lys	Thr																														
																		275													280													285	
Lys	Ala	Phe	Ile	Thr	His	Gly	Gly	Met	Asn	Gly	Ile	Tyr	Glu	Ala	Ile																														
																		290													295													300	
Tyr	His	Gly	Val	Pro	Met	Val	Gly	Val	Pro	Ile	Phe	Gly	Asp	Gln	Leu																														
305																			310													315													320
Asp	Asn	Ile	Ala	His	Met	Lys	Ala	Lys	Gly	Ala	Ala	Val	Glu	Ile	Asn																														
																		325													330													335	
Phe	Lys	Thr	Met	Thr	Ser	Glu	Asp	Leu	Leu	Arg	Ala	Leu	Arg	Thr	Val																														
																		340													345													350	
Ile	Thr	Asp	Ser	Ser	Tyr	Lys	Glu	Asn	Ala	Met	Arg	Leu	Ser	Arg	Ile																														
																		355													360													365	
His	His	Asp	Gln	Pro	Val	Lys	Pro	Leu	Asp	Arg	Ala	Val	Phe	Trp	Ile																														
																		370													375													380	
Glu	Phe	Val	Met	Arg	His	Lys	Gly	Ala	Lys	His	Leu	Arg	Ser	Ala	Ala																														
385																			390													395													400
His	Asp	Leu	Thr	Trp	Phe	Gln	His	Tyr	Ser																																				

1. The first group of people who are interested in the study of the history of the United States are the people who are interested in the history of the United States. This group of people is interested in the history of the United States because they want to know more about the United States. They want to know more about the United States because they want to know more about the United States.

Met	Ala	Pro	Gly	Lys	Leu	Ala	Ser	Ala	Val	Leu	Leu	Leu	Leu	Leu	Cys
1				5				10						15	
Cys	Ala	Gly	Ser	Gly	Phe	Cys	Gly	Lys	Val	Leu	Val	Trp	Pro	Cys	Glu
			20					25					30		
Met	Ser	His	Trp	Leu	Asn	Leu	Lys	Thr	Leu	Leu	Glu	Glu	Leu	Val	Lys
		35					40					45			
Arg	Gly	His	Glu	Val	Thr	Val	Leu	Thr	Leu	Ser	Asn	Asn	Leu	Phe	Ile
	50					55					60				
Asp	Tyr	Asn	Arg	His	Pro	Ala	Phe	Asn	Phe	Glu	Val	Ile	Pro	Val	Pro
65					70					75					80
Thr	Asp	Lys	Asn	Met	Ser	Glu	Asn	Ile	Leu	Asn	Glu	Phe	Ile	Glu	Leu
				85					90					95	
Ala	Val	Asn	Val	Met	Pro	Thr	Met	Pro	Leu	Trp	Gln	Ser	Gly	Lys	Leu
			100					105					110		
Leu	Gln	Gln	Phe	Phe	Val	Gln	Ile	Thr	Glu	Asp	Leu	Gly	Leu	Asn	Cys
		115					120					125			
Arg	Asn	Thr	Val	Tyr	Asn	Gln	Ser	Leu	Met	Lys	Lys	Leu	Arg	Asp	Ser
	130					135					140				
Lys	Tyr	Asp	Val	Leu	Val	Thr	Asp	Pro	Val	Ile	Pro	Cys	Gly	Glu	Leu
145					150					155					160
Val	Ala	Glu	Met	Leu	Gly	Val	Pro	Phe	Val	Asn	Met	Leu	Lys	Phe	Ser
				165					170					175	
Met	Gly	His	Thr	Ile	Glu	Lys	Tyr	Cys	Gly	Gln	Leu	Pro	Ala	Pro	Pro
			180					185					190		
Ser	Tyr	Val	Pro	Val	Pro	Leu	Gly	Gly	Leu	Thr	Thr	Arg	Met	Thr	Phe
	195						200					205			
Met	Glu	Arg	Val	Lys	Asn	Met	Val	Phe	Ser	Val	Leu	Phe	Asp	Phe	Trp
	210				215						220				
Ile	Gln	Gln	Tyr	Asp	Tyr	Lys	Phe	Trp	Asp	Gln	Phe	Tyr	Ser	Glu	Ala
225					230					235					240
Leu	Gly	Arg	Pro	Thr	Thr	Leu	Cys	Glu	Ile	Met	Gly	Lys	Ala	Glu	Ile
				245					250					255	
Trp	Leu	Ile	Arg	Thr	Tyr	Trp	Asp	Phe	Glu	Phe	Pro	Arg	Pro	Tyr	Leu
			260					265					270		
Pro	Asn	Phe	Glu	Phe	Val	Gly	Gly	Leu	His	Cys	Lys	Pro	Ala	Lys	Pro
		275					280					285			
Leu	Pro	Lys	Glu	Met	Glu	Glu	Phe	Val	Gln	Ser	Ser	Gly	Glu	Asp	Gly
	290					295					300				
Val	Val	Val	Phe	Ser	Leu	Gly	Ser	Met	Val	Lys	Asn	Leu	Thr	Glu	Glu
305					310					315					320
Lys	Ala	Asn	Leu	Ile	Ala	Ser	Ala	Leu	Ala	Gln	Ile	Pro	Gln	Lys	Val
				325					330					335	
Leu	Trp	Arg	Tyr	Lys	Gly	Lys	Lys	Pro	Ala	Thr	Leu	Gly	Pro	Asn	Thr
			340					345					350		
Arg	Leu	Phe	Asp	Trp	Ile</										

385					390					395				400
Pro	Asp	Asn	Leu	Ala	Gly	Met	Lys	Ala	Lys	Gly	Ala	Ala	Val	Glu
				405					410					415
Asn	Met	Asn	Thr	Met	Thr	Ser	Ala	Asp	Leu	Leu	Gly	Ala	Leu	Arg
			420					425					430	
Val	Ile	Asn	Asp	Pro	Thr	Tyr	Lys	Glu	Asn	Ala	Met	Lys	Leu	Ser
		435					440					445		
Ile	His	His	Asp	Gln	Pro	Val	Lys	Pro	Leu	Asp	Arg	Ala	Ala	Phe
	450					455					460			
Val	Glu	Phe	Val	Met	His	His	Lys	Gly	Ala	Lys	His	Leu	Arg	Val
465					470					475				480
Ala	His	Asp	Leu	Ser	Trp	Phe	Gln	Tyr	His	Ser	Leu	Asp	Val	Ile
			485						490					495
Phe	Leu	Leu	Ala	Cys	Val	Ala	Ser	Ala	Ile	Leu	Leu	Val	Thr	Lys
			500					505					510	
Cys	Leu	Phe	Ser	Phe	Gln	Asn	Phe	Ile	Lys	Ile	Gly	Lys	Arg	Ile
		515					520					525		
Lys	Glu													
	530													

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000